

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,840

DATE: 04/04/2001

TIME: 11:38:37

Input Set : A:\5787577.app

Output Set: N:\CRF3\04042001\I720840.raw

ENTERED

3 <110> APPLICANT: Biotica Technology Limited
 4 Leadlay, Peter F
 5 Staunton, James
 6 Cortes, Jesus
 7 McArthur, Hamish AI
 9 <120> TITLE OF INVENTION: Polyketides and their synthesis
 11 <130> FILE REFERENCE: IS/CP5787577
 C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/720,840
 C--> 14 <141> CURRENT FILING DATE: 2000-12-29
 16 <150> PRIOR APPLICATION NUMBER: GB 9814006.4
 17 <151> PRIOR FILING DATE: 1998-06-29
 19 <160> NUMBER OF SEQ ID NOS: 55
 21 <170> SOFTWARE: PatentIn Ver. 2.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 398
 25 <212> TYPE: PRT
 26 <213> ORGANISM: Streptomyces sp. C5
 28 <400> SEQUENCE: 1
 29 Met Val Thr Gly Leu Gly Ile Val Ala Pro Asn Gly Leu Gly Val Gly
 30 1 5 10 15
 32 Ala Ile Trp Asp Ala Val Leu Asn Gly Arg Asn Gly Ile Gly Pro Leu
 33 20 25 30
 35 Arg Arg Phe Ala Asp Asp Gly Arg Leu Gly Arg Leu Ala Gly Glu Val
 36 35 40 45
 38 Ser Asp Phe Val Pro Glu Asp His Leu Pro Lys Arg Leu Leu Val Gln
 39 50 55 60
 41 Thr Asp Pro Met Thr Gln Met Thr Ala Leu Ala Ala Ala Glu Trp Ala
 42 65 70 75 80
 44 Leu Arg Glu Ala Gly Cys Ala Pro Ser Ser Pro Leu Glu Ala Gly Val
 45 85 90 95
 47 Ile Thr Ala Ser Ala Ser Gly Gly Phe Ala Ser Gly Gln Arg Glu Leu
 48 100 105 110
 50 Gln Asn Leu Trp Ser Lys Gly Pro Ala His Val Ser Ala Tyr Met Ser
 51 115 120 125
 53 Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile Ala Ile Arg His
 54 130 135 140
 56 Asp Leu Arg Gly Pro Val Gly Val Val Val Ala Glu Gln Ala Gly Gly
 57 145 150 155 160
 59 Leu Asp Ala Leu Ala His Ala Arg Arg Lys Val Arg Gly Gly Ala Glu
 60 165 170 175
 62 Leu Ile Val Ser Gly Ala Met Asp Ser Ser Leu Cys Pro Tyr Gly Met
 63 180 185 190
 65 Ala Ala Gln Val Arg Ser Gly Arg Leu Ser Gly Ser Asp Asp Pro Thr
 66 195 200 205
 68 Ala Gly Tyr Leu Pro Phe Asp Arg Arg Ala Ala Gly His Val Pro Gly
 69 210 215 220
 71 Glu Gly Gly Ala Ile Leu Ala Val Glu Asp Ala Glu Arg Val Ala Glu

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,840

DATE: 04/04/2001

TIME: 11:38:37

Input Set : A:\5787577.app

Output Set: N:\CRF3\04042001\I720840.raw

```

72 225                230                235                240
74 Arg Gly Gly Lys Val Tyr Gly Ser Ile Ala Gly Thr Ala Ser Phe Asp
75                245                250                255
77 Pro Pro Pro Gly Ser Gly Arg Pro Ser Ala Leu Ala Arg Ala Val Glu
78                260                265                270
80 Thr Ala Leu Ala Asp Ala Gly Leu Asp Arg Ser Asp Ile Ala Val Val
81                275                280                285
83 Phe Ala Asp Gly Ala Ala Val Gly Glu Leu Asp Val Ala Glu Ala Glu
84                290                295                300
86 Ala Leu Ala Ser Val Phe Gly Pro His Arg Val Pro Val Thr Val Pro
87 305                310                315                320
89 Lys Thr Leu Thr Gly Arg Leu Tyr Ser Gly Ala Gly Pro Leu Asp Val
90                325                330                335
92 Ala Thr Gly Leu Leu Ala Leu Arg Asp Glu Val Val Pro Ala Thr Gly
93                340                345                350
95 His Val His Pro Asp Pro Asp Leu Pro Leu Asp Val Val Thr Gly Arg
96                355                360                365
98 Pro Arg Ala Met Ala Asp Ala Arg Ala Ala Leu Val Val Ala Arg Gly
99                370                375                380
101 His Gly Gly Phe Asn Ser Ala Leu Val Val Arg Gly Ala Ala
102 385                390                395
106 <210> SEQ ID NO: 2
107 <211> LENGTH: 425
108 <212> TYPE: PRT
109 <213> ORGANISM: Streptomyces peucetius
111 <400> SEQUENCE: 2
112 Met Thr Gly Thr Ala Ala Arg Thr Ala Ser Ser Gln Leu His Ala Ser
113 1 5 10 15
115 Pro Ala Gly Arg Arg Gly Leu Arg Gly Arg Ala Val Val Thr Gly Leu
116 20 25 30
118 Gly Ile Val Ala Pro Asn Gly Leu Gly Val Gly Ala Tyr Trp Asp Ala
119 35 40 45
121 Val Leu Asn Gly Arg Asn Gly Ile Gly Pro Leu Arg Arg Phe Thr Gly
122 50 55 60
124 Asp Gly Arg Leu Gly Arg Leu Ala Gly Glu Val Ser Asp Phe Val Pro
125 65 70 75 80
127 Glu Asp His Leu Pro Lys Arg Leu Leu Ala Gln Thr Asp Pro Met Thr
128 85 90 95
130 Gln Tyr Ala Leu Ala Ala Ala Glu Trp Ala Leu Arg Glu Ser Gly Cys
131 100 105 110
133 Ser Pro Ser Ser Pro Leu Glu Ala Gly Val Ile Thr Ala Ser Ala Ser
134 115 120 125
136 Gly Gly Phe Ala Phe Gly Gln Arg Glu Leu Gln Asn Leu Trp Ser Lys
137 130 135 140
139 Gly Pro Ala His Val Ser Ala Tyr Met Ser Phe Ala Trp Phe Tyr Ala
140 145 150 155 160
142 Val Asn Thr Gly Gln Ile Ala Ile Arg His Asp Leu Arg Gly Pro Val
143 165 170 175
145 Gly Val Val Val Ala Glu Gln Ala Gly Gly Leu Asp Ala Leu Ala His

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,840

DATE: 04/04/2001

TIME: 11:38:37

Input Set : A:\5787577.app

Output Set: N:\CRF3\04042001\I720840.raw

```

146          180          185          190
148 Ala Arg Arg Lys Val Arg Gly Gly Ala Glu Leu Ile Val Ser Gly Ala
149          195          200          205
151 Val Asp Ser Ser Leu Cys Pro Tyr Gly Met Ala Ala Gln Val Lys Ser
152          210          215          220
154 Gly Arg Leu Ser Gly Ser Asp Asn Pro Thr Ala Gly Tyr Leu Pro Phe
155 225          230          235          240
157 Asp Arg Arg Ala Ala Gly His Val Pro Gly Glu Gly Gly Ala Ile Leu
158          245          250          255
160 Thr Val Glu Asp Ala Glu Arg Ala Ala Glu Arg Gly Ala Lys Val Tyr
161          260          265          270
163 Gly Ser Ile Ala Gly Tyr Gly Ala Ser Phe Asp Pro Pro Gly Ser
164          275          280          285
166 Gly Arg Pro Ser Ala Leu Ala Arg Ala Val Glu Thr Ala Leu Ala Asp
167          290          295          300
169 Ala Gly Leu Asp Gly Ser Asp Ile Ala Val Val Phe Ala Asp Gly Ala
170 305          310          315          320
172 Ala Val Pro Glu Leu Asp Ala Ala Glu Ala Glu Ala Leu Ala Ser Val
173          325          330          335
175 Phe Gly Pro Arg Arg Val Pro Val Thr Val Pro Lys Thr Leu Thr Gly
176          340          345          350
178 Arg Leu Tyr Ser Gly Ala Gly Pro Leu Asp Val Ala Thr Ala Leu Leu
179          355          360          365
181 Ala Leu Arg Asp Glu Val Val Pro Ala Thr Ala His Val Asp Pro Asp
182          370          375          380
184 Pro Asp Leu Pro Leu Asp Val Val Thr Gly Arg Pro Arg Ser Leu Ala
185 385          390          395          400
187 Asp Ala Arg Ala Ala Leu Leu Val Ala Arg Gly Tyr Gly Gly Phe Asn
188          405          410          415
190 Ser Ala Leu Val Val Arg Gly Ala Ala
191          420          425
195 <210> SEQ ID NO: 3
196 <211> LENGTH: 407
197 <212> TYPE: PRT
198 <213> ORGANISM: Streptomyces coelicolor
200 <400> SEQUENCE: 3
201 Met Ser Val Leu Ile Thr Gly Val Gly Val Val Ala Pro Asn Gly Leu
202 1          5          10          15
204 Gly Leu Ala Pro Tyr Trp Ser Ala Val Leu Asp Gly Arg His Gly Leu
205          20          25          30
207 Gly Pro Val Thr Arg Phe Asp Val Ser Arg Tyr Pro Ala Thr Leu Ala
208          35          40          45
210 Gly Gln Ile Asp Asp Phe His Ala Pro Asp His Ile Pro Gly Arg Leu
211          50          55          60
213 Leu Pro Gln Thr Asp Pro Ser Thr Arg Leu Ala Leu Thr Ala Ala Asp
214          65          70          75          80
216 Trp Ala Leu Gln Asp Ala Lys Ala Asp Pro Glu Ser Leu Thr Asp Tyr
217          85          90          95
219 Asp Met Gly Val Val Thr Ala Asn Ala Cys Gly Gly Phe Asp Phe Thr

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,840

DATE: 04/04/2001

TIME: 11:38:37

Input Set : A:\5787577.app

Output Set: N:\CRF3\04042001\I720840.raw

```

220          100          105          110
222 His Arg Glu Phe Arg Lys Leu Trp Ser Glu Gly Pro Lys Ser Val Ser
223          115          120          125
225 Val Tyr Glu Ser Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile
226          130          135          140
228 Ser Ile Arg His Gly Met Arg Gly Pro Ser Ser Ala Leu Val Ala Glu
229 145          150          155          160
231 Gln Ala Gly Gly Leu Asp Ala Leu Gly His Ala Arg Arg Thr Ile Arg
232          165          170          175
234 Arg Gly Thr Pro Leu Val Val Ser Gly Gly Val Asp Ser Ala Leu Asp
235          180          185          190
237 Pro Trp Gly Trp Val Ser Gln Ile Ala Ser Gly Arg Ile Ser Thr Ala
238          195          200          205
240 Thr Asp Pro Asp Arg Ala Tyr Leu Pro Phe Asp Glu Arg Ala Ala Gly
241          210          215          220
243 Tyr Val Pro Gly Glu Gly Gly Ala Ile Leu Val Leu Glu Asp Ser Ala
244 225          230          235          240
246 Ala Ala Glu Ala Arg Gly Arg His Asp Ala Tyr Gly Glu Leu Ala Gly
247          245          250          255
249 Cys Ala Ser Thr Phe Asp Pro Ala Pro Gly Ser Gly Arg Pro Ala Gly
250          260          265          270
252 Leu Glu Arg Ala Ile Arg Leu Ala Leu Asn Asp Ala Gly Thr Gly Pro
253          275          280          285
255 Glu Asp Val Asp Val Val Phe Ala Asp Gly Ala Gly Val Pro Glu Leu
256          290          295          300
258 Asp Ala Ala Glu Ala Arg Ala Ile Gly Arg Val Phe Gly Arg Glu Gly
259 305          310          315          320
261 Val Pro Val Thr Val Pro Lys Thr Thr Thr Gly Arg Leu Tyr Ser Gly
262          325          330          335
264 Gly Gly Pro Leu Asp Val Val Thr Ala Leu Met Ser Leu Arg Glu Gly
265          340          345          350
267 Val Ile Ala Pro Thr Ala Gly Val Thr Ser Val Pro Arg Glu Tyr Gly
268          355          360          365
270 Ile Asp Leu Val Leu Gly Glu Pro Arg Ser Thr Ala Pro Arg Thr Ala
271          370          375          380
273 Leu Val Leu Ala Arg Gly Arg Trp Gly Phe Asn Ser Ala Ala Val Leu
274 385          390          395          400
276 Arg Arg Phe Ala Pro Thr Pro
277          405
281 <210> SEQ ID NO: 4
282 <211> LENGTH: 403
283 <212> TYPE: PRT
284 <213> ORGANISM: Saccharopolyspora hirsuta
286 <400> SEQUENCE: 4
287 Met Ser Thr Trp Val Thr Gly Met Gly Val Val Ala Pro Asn Gly Leu
288 1          5          10          15
290 Gly Ala Asp Asp His Trp Ala Ala Thr Leu Lys Gly Arg His Gly Ile
291          20          25          30
293 Ser Arg Leu Ser Arg Phe Asp Pro Thr Gly Tyr Pro Ala Glu Leu Ala

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,840

DATE: 04/04/2001

TIME: 11:38:37

Input Set : A:\5787577.app

Output Set: N:\CRF3\04042001\I720840.raw

```

294          35          40          45
296 Gly Gln Val Leu Asp Phe Asp Ala Thr Glu His Leu Pro Lys Arg Leu
297          50          55          60
299 Leu Pro Gln Thr Asp Val Ser Thr Arg Phe Ala Leu Ala Ala Ala Ala
300 65          70          75          80
302 Trp Ala Leu Ala Asp Ala Glu Val Asp Pro Ala Glu Leu Pro Glu Tyr
303          85          90          95
305 Gly Thr Gly Val Ile Thr Ser Asn Ala Thr Gly Gly Phe Glu Phe Thr
306          100          105          110
308 His Arg Glu Phe Arg Lys Leu Trp Ala Gln Gly Pro Glu Phe Val Ser
309          115          120          125
311 Val Tyr Glu Ser Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile
312          130          135          140
314 Ser Ile Arg His Gly Leu Arg Gly Pro Gly Ser Val Leu Val Ala Glu
315 145          150          155          160
317 Gln Ala Gly Gly Leu Asp Ala Val Gly His Gly Gly Ala Val Arg Asn
318          165          170          175
320 Gly Thr Pro Met Val Val Thr Gly Gly Val Asp Ser Ser Phe Asp Pro
321          180          185          190
323 Trp Gly Trp Val Ser His Val Ser Ser Gly Arg Val Ser Arg Ala Thr
324          195          200          205
326 Asp Pro Gly Arg Ala Tyr Leu Pro Phe Asp Val Ala Ala Asn Gly Tyr
327          210          215          220
329 Val Pro Gly Glu Gly Gly Ala Ile Leu Leu Leu Glu Asp Ala Glu Ser
330 225          230          235          240
332 Ala Lys Ala Arg Gly Ala Thr Gly Tyr Gly Glu Ile Ala Gly Tyr Ala
333          245          250          255
335 Ala Thr Phe Asp Pro Ala Pro Gly Ser Glu Arg Pro Pro Ala Leu Arg
336          260          265          270
338 Arg Ala Ile Glu Leu Ala Leu Ala Asp Ala Glu Leu Arg Pro Glu Gln
339          275          280          285
341 Val Asp Val Val Phe Ala Asp Ala Ala Gly Val Ala Glu Leu Asp Ala
342          290          295          300
344 Ile Glu Ala Ala Ala Ile Arg Glu Leu Phe Gly Pro Ser Gly Val Pro
345 305          310          315          320
347 Val Thr Ala Pro Lys Thr Met Thr Gly Arg Leu Tyr Ser Gly Gly Gly
348          325          330          335
350 Pro Leu Asp Leu Val Ala Ala Leu Leu Ala Ile Arg Asp Gly Val Ile
351          340          345          350
353 Pro Pro Thr Val His Thr Ala Glu Pro Val Pro Glu His Gln Leu Asp
354          355          360          365
356 Leu Val Thr Gly Asp Pro Arg His Gln Gln Leu Gly Thr Ala Leu Val
357          370          375          380
359 Leu Ala Arg Gly Lys Trp Gly Phe Asn Ser Ala Val Val Val Arg Gly
360 385          390          395          400
362 Val Thr Gly
367 <210> SEQ ID NO: 5
368 <211> LENGTH: 415
369 <212> TYPE: PRT

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/720,840

DATE: 04/04/2001

TIME: 11:38:38

Input Set : A:\5787577.app

Output Set: N:\CRF3\04042001\I720840.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application Number

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date